INTRODUCTORY VERSION

Database Systems

An Application-Oriented Approach

SECOND EDITION

Michael **KIFER** Arthur **BERNSTEIN** Philip M. **LEWIS**

INTRODUCTORY VERSION

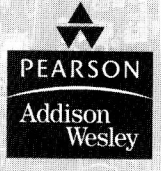
Database Systems

An Application-Oriented Approach

SECOND EDITION

Michael Kifer Arthur Bernstein Philip M. Lewis

STATE UNIVERSITY OF NEW YORK, STONY BROOK



Boston San Francisco New York
London Toronto Sydney Tokyo Singapore Madrid
Mexico City Munich Paris Cape Town Hong Kong Montreal

Executive Editor	Susan Hartman Sullivan
Acquisitions Editor	Matt Goldstein
Project Editor	Maite Suarez-Rivas
Marketing Manager	Nathan Schultz
Senior Marketing Coordinator	Lesly Hershman
Senior Production Supervisor	Jeffrey Holcomb
Project Management	Windfall Software
Copyeditor	Elisabeth Beller
Proofreader	Jennifer McClain
Composition	Windfall Software, using ZzTEX
Text Designer	Sandra Rigney
Cover Designer	Joyce Cosentino Wells
Cover Photo	© 2004 Digital Vision
Prepress and Manufacturing	Caroline Fell

Access the latest information about Addison-Wesley titles from our World Wide Web site: http://www.aw-bc.com/computing

Library of Congress Cataloging-in-Publication Data

Kifer, M. (Michael), 1954–
Database systems : an application-oriented approach / Michael Kifer, Arthur Bernstein, Philip M. Lewis.—Introductory version, 2nd ed.
p. cm.
Lewis's name appears first on the earlier edition.
Rev. ed. of: Databases and transaction processing / Philip M. Lewis. 2002.
ISBN 0-321-22838-3 (hardcover)
1. Database management. 2. Transaction systems (Computer systems) I. Bernstein, Arthur J. II. Lewis, Philip M., 1931– Databases and transaction processing. III. Title.

QA76.9.D3K4965 2004
005.74—dc22 2003068925

ISBN 0-321-22838-3
1 2 3 4 5 6 7 8 9 10—CRW—08 07 06 05

INTRODUCTORY VERSION

Database Systems

An Application-Oriented Approach

SECOND EDITION

In memory of my late parents, Luba and Isaac;
and to my wife, Lora, and my children. M.K.

To my wife, Edie, my children, and my grandchildren. A.J.B.

To my wife, Rhoda, my children, and my grandchildren. P.M.L.

Preface

We are publishing the second edition of our textbook in two versions:

- This version, which consists of introductory material, is appropriate for a first undergraduate or graduate course in databases.
- The second version, which is the complete book, is appropriate for three courses:
  - An introductory undergraduate or graduate course in databases
  - An undergraduate or graduate course in transaction processing for students who have had an introductory course in databases
  - An advanced undergraduate or a first graduate course in databases for students who have had an introductory course in databases

One of our goals was to reduce the size and make this introductory version more affordable to students. Another was to capitalize on our experience in using the first edition of the book to make an even better introductory text.

The chapters in this book are not just a subset of those in the complete book. We believe that instructors of an introductory database course should have the option of enriching an introductory course by including material on object databases and XML—topics that are covered in great detail in several chapters in the complete book. Therefore we have added to the introductory book two new chapters, Chapter 16, Introduction to Object Databases, and Chapter 17, Introduction to XML and Web Data, which contain an appropriately chosen subset of the material in the full version of this book.

To keep the book up-to-date with the rapidly changing technology, we have added a substantial amount of material on UML to a number of chapters and have included a new chapter on Database Tuning, Chapter 12, in both the introductory and complete books.

As with the first edition, our focus is on how to build applications using databases rather than on how to build the database management system itself. We believe that many more students will be implementing applications than will be building DBMSs. Thus, we include substantial material describing the languages and APIs used by transactions to access a database, such as embedded SQL, ODBC, and JDBC.

Although we cover many practical aspects of database and transaction processing applications, we are primarily concerned with the concepts that underlie these topics rather than with the details of particular commercial systems or applications.

Thus we concentrate on the concepts behind the relational and object data models. These concepts will remain the foundation of database processing long after SQL is obsolete.

To enhance students' understanding of the technical material, we have included a case study of a transaction processing application, the Student Registration System, which is carried through the book. While a student registration system can hardly be considered glamorous, it has the unique advantage that all students have interacted with such a system as users. More importantly, it turns out to be a surprisingly rich application, so we can use it to illustrate many of the issues in database design, query processing, and transaction processing.

A unique aspect of the book is a presentation of the software engineering concepts required to implement transaction processing applications, using the Student Registration System as an example. Since the implementations of many information systems fail because of poor project management and inadequate software engineering, we feel that these topics should be an important part of the student's education. Our treatment of software engineering issues is brief, since many students will take a separate course in this subject. However, we believe that they will be better able to understand and apply that material when they see it presented in the context of an information system implementation. Since the courses that use this text at Stony Brook are not software engineering courses, we do not cover this material in class. Instead, we ask the students to read it and require that they use good software engineering practice in their class projects. We do cover in class those aspects of the Student Registration System that illustrate important issues in databases and transaction processing.

Changes in the Second Edition

The technology underlying database and transaction processing systems is changing so rapidly that we have made a large number of changes and additions to the material of the first edition. One rapidly advancing technology is the Unified Modeling Language, UML. We added substantial amount of material on UML in Chapter 4 on database design, in addition to the material on E-R diagrams that was already there. We also added UML to the material on software engineering in Chapters 2, 14, and 15.

A new chapter on Database Tuning, Chapter 12, was added because so much effort in the real world is spent increasing the throughput of database and transaction processing applications.

In addition, material has been added and updated in almost all the chapters. Significant examples of this are the coverage of SQL/XML and RAID technology.

One important area that is *not* included in this volume is Web Services. Since this is a rapidly developing and interesting application-oriented subject we have significantly revised the compete version of this text to include material on this topic. In addition to strengthening the book on the subject of XML Technology by updating the chapter on XML and Web Data and adding a section on SQL/XML, we have added a new chapter on Web Services that contains material on SOAP,

WSDL, BPEL, UDDI, and XML-based transaction processing using WS-Coordination and WS-Transaction. In the chapter on Security and Internet Commerce, we added a section on XML-based encryption, using XML-Encryption, XML-Signature, WS-Security, and SAML. And in the chapter on Architecture of Transaction Processing Systems, we added material on Web Application Servers and J2EE, which are used to implement the back-end of many Web services.

Organization of the Book

Chapters 1 through 7 should be taught in the order in which they appear in the book. Chapter 8 contains much of the information that students need in order to put the knowledge they acquired in the preceding chapters into practice. However, subsequent chapters do not significantly depend on Chapter 8. Chapters 9 through 12 in Part 3 should be taught sequentially. Chapter 13 in the same part is largely independent. The software engineering chapters in Part 4 utilize the material of the chapters in Parts 2 and 3, but the software engineering chapters can be read in parallel with the database material. Chapters 16 and 17 in the advanced part of the book depend on the first seven chapters in Part 2.

Finally we note that the sections in this book that are marked with an asterisk (*) are optional and can be omitted, if the instructor prefers to do so. Sections marked with the Ⓒ icon in the table of contents deal with the case study. Also, exercises that are marked with an asterisk are slightly harder than the rest, and exercises that are marked with two asterisks are even harder.

Supplements

In addition to the text, the following supplementary materials are available to assist instructors:

- Online PowerPoint presentations for all chapters
- Online PowerPoint slides of all figures
- An online solution manual containing solutions for the exercises
- Additional references, notes, errata, homeworks, and exams.

For more information on obtaining these supplements, please visit this book's Companion Website at *www.aw-bc.com/kifer*. The solutions manual and PowerPoint presentations are available only to instructors through your Addison-Wesley sales representative. To contact your representative, please visit *www.aw.com*.

Acknowledgments

We would like to thank the reviewers, whose comments and suggestions significantly improved the second edition of the book:

Tran Cao Son, New Mexico State University

Frantisek Franek, McMaster University

Junping Sun, Nova Southeastern University
Philip Cannata, Sun Microsystems
Dehu Qi, Lamar University
Nematollah Shiri, Concordia University
Jian Pei, State University of New York at Buffalo
Jack Wileden, University of Massachusetts Amherst
Sibel Adali, Rensselaer Polytechnic Institute
Roger King, University of Colorado at Boulder
Markus Schneider, University of Florida
Yaron Y. Goland, BEA
Dennis Shasha, New York University
Christelle Scharff, Pace University
Zhiwei Wang, Graduate Programs in Software Engineering, IT, and IS University of St. Thomas.

We would also like to thank the reviewers of the first edition of the book:

Suad Alagic, Wichita University
Catriel Beeri, The Hebrew University
Rick Cattel, Sun Microsystems
Jan Chomicki, SUNY Buffalo
Henry A. Etlinger, Rochester Institute of Technology
Leonidas Fegaras, University of Texas at Arlington
Alan Fekete, University of Sidney
Johannes Gehrke, Cornell University
Hershel Gottesman, Consultant
Jiawei Han, Simon Fraser University
Peter Honeyman, University of Michigan
Vijay Kumar, University of Missouri–Kansas City
Jonathan Lazar, Towson University
Dennis McLeod, University of Southern California
Rokia Missaoui, University of Quebec in Montreal
Clifford Neuman, University of Southern California
Fabian Pascal, Consultant
Sudha Ram, University of Arizona
Krithi Ramamritham, University of Massachusetts–Amherst, and IIT Bombay
Andreas Reuter, International University in Germany, Bruchsal
Arijit Sengupta, Georgia State University

Munindar P. Singh, North Carolina State University

Greg Speegle, Baylor University

Junping Sun, Nova Southeastern University

Joe Trubisz, Consultant

Vassilis J. Tsotras, University of California, Riverside

Emilia E. Villarreal, California Polytechnic State University

We would also like to thank the following people who were kind enough to provide us with additional information and answers to our questions: Don Chamberlin, Daniela Florescu, Jim Gray, Pankaj Gupta, Rob Kelly, and C. Mohan.

Two people taught out of beta versions of the book and made useful comments and suggestions: David S. Warren and Radu Grosu. Joe Trubicz served not only as a reviewer when the manuscript was complete, but provided critical comments on early versions of many of the chapters.

A number of students were very helpful in reading and checking the correctness of various parts of the book: Ziyang Duan, Shiyong Lu, Swapnil Patil, Guizhen Yang, and Yan Zhang.

Many thanks to the staff of the Computer Science Department at Stony Brook, and in particular Kathy Germana, who helped make things happen at work.

We would particularly like to thank Matt Goldstein and Maite Suarez-Rivas, our editors at Addison-Wesley, who played an important role in shaping the contents and approach of the book in its early stages and throughout the time we were writing it. We would also like to thank the various staff members of Addison-Wesley and Windfall Software, who did an excellent job of editing and producing the book: Jeffrey Holcomb, Paul Anagnostopoulos, Elisabeth Beller, Jennifer McClain, and Joe Snowden.

Last, but not least, we would like to thank our wives, Lora, Edie, and Rhoda, who provided much needed support and encouragement while we were writing the book.

Contents